

Paper overview

ATLAS Draft

Measurements of underlying event properties using calorimeter clusters in pp collisions at 900 GeV and 7 TeV with the ATLAS detector at the LHC

Version: 1.0

To be submitted to: EPJC

Corresponding editor(s)

S. Chekanov (chakanau@hep.anl.gov)
J. Proudfoot (Proudfoot@anl.gov)

A list of supporting internal notes and their authors can be found at:

<https://wiki.cern.ch/wiki/bin/view/AtlasProtected/TopoClustersUE>

Supporting internal notes

ATL-COM-PHYS-2010-210	http://cdsweb.cern.ch/record/1262602
ATL-COM-PHYS-2010-293	http://cdsweb.cern.ch/record/1267398
ATL-COM-PHYS-2010-440	http://cdsweb.cern.ch/record/1274248
ATL-COM-PHYS-2010-351	http://cdsweb.cern.ch/record/1271366
ATL-COM-PHYS-2010-165	http://cdsweb.cern.ch/record/1256572
ATL-COM-PHYS-2010-367	http://cdsweb.cern.ch/record/1271712

Editorial Board

Donatella Cavalli (chair)	(Donatella.Cavalli@mi.infn.it)
William Bell	(W.Bell@cern.ch)
Ana Henriques	(Ana.Maria.Henriques.Correia@cern.ch)
Heiko Lacker	(lacker@physik.hu-berlin.de)

Comments are due by: XXXX, 2011



document created on January 22, 2011 from file ueclusters_paper.tex
cover page automatically created with atlascover.sty

+ J.Butterworth, K. Einsweiler, E.Nurse

Web page

<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopoClustersUE>

TopoClustersUE

- ↓ [Studies of particle flow using calorimeter clusters in pp collisions at 900 GeV and 7 TeV with the ATLAS detector at the LHC](#)
- ↓ [People](#)
- ↓ [Paper outline](#)
- ↓ [Introduction](#)
- ↓ [Supporting material](#)
 - ↓ [Notes](#)
 - ↓ [Talks](#)
 - ↓ [Histograms and figures](#)
- ↓ [Proposed final plots](#)
- ↓ [Analysis Code](#)
 - ↓ [Topocluster energy scale](#)

Contains drafts, talks, paper outline, CONF draft

Studies of particle flow using calorimeter clusters in pp collisions at 900 GeV and 7 TeV with the ATLAS detector at the LHC

People

C. Bertella , A. Buckley, S. Chekanov , P. Giovaninni , N. Kanaya ,D. Kar, A. Moraes, S. Menke , J. Nielsen, G.A. Hare, J. Proudfoot , C. Roda , P.Starovoitov, I. Vivarelli , R. Yoshida , J. Zhang

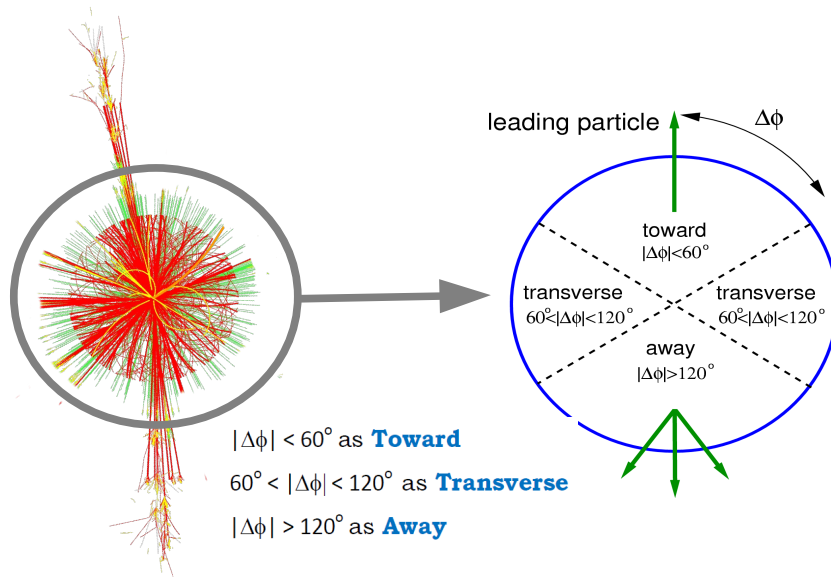
+ Emily Nurse (softQCD coordinator)

Editorial board

- Donatella Cavalli (chair)
- Will Bell
- Ana Henriques
- Heiko Lacker



Studies of Underlying Event. Definitions.



- **Soft QCD is unsolved from the first principles**
- **Experimentally:**
 - Isolate softQCD particle activity by looking at the “transverse” region
- **Look at models**
 - tune them to control softQCD activity for high-precision measurements

Commonly used variables:

- p_T^{lead} - Transverse momentum of the stable particle with maximum p_T in the event (approximates the direction of hard interaction in MinBias data)
- $\langle d^2 N / d\eta d\phi \rangle$ - Mean number of stable particles per unit η - ϕ
- $\langle d^2 \sum p_T / d\eta d\phi \rangle$ - Mean scalar p_T sum of stable particles per η - ϕ (independent of density)

+ more ...

ATLAS arXiv:1012.0791

“*stable particle*” → stable charged particle for tracking analysis



Event & Topocluster selection: 900 GeV

Good run list for: 141565-142383

- MinBias Monte Carlo sample: ATLAS-GEO-08-00-02 (+ more updated geometries as a check)
- L1_MBTS_1 trigger. Good primary vertex
- **Selection of topoclusters:**
 - Topoclusters after local **hadronic calibration** (EM-scale as systematics checks)
 - Concentrate on the central region $|\eta| < 2.5$ (easy cross check with tracks)
 - $p_T > 500$ MeV (as for the tracking analysis)
 - leading cell energy of the cluster is required to be less than 90% of the cluster energy;
 - the energy sampling maximum should not be in a calorimeter region without good calibration;
 - the fraction of energy associated with bad cells should be less than 50%

Event & Topocluster selection: 7 TeV

- ◆ Good run & lumi blocks for 152166- 152844. Lumi $\sim 238 \mu\text{b}^{-1}$ (about 7M events)
- ◆ MinBias Monte Carlo sample: ATLAS-GEO-10-00-00
- ◆ Same event cuts as for 900 GeV + pile-up removal
- ◆ Same cuts on topoclusters

Used primary tracks to re-weight MC topocluster distributions to match the data

- track selection as for the MinBias UE paper



Monte Carlo models

- **PYTHIA 6**, actually 6.4.21: pT-ordered parton shower, MRST LO PDF, multiple parton-parton scattering, string fragmentation
- **PYTHIA ATLAS MC09**: parameters tuned to underlying events and minimum bias data from Tevatron at 630 GeV to 1.8 TeV (ATLAS optimization)
- **PYTHIA ATLAS MC09c**: MC09 optimizing the strength of the color reconnection to describe pT dependence on N(ch) in the CDF data at 1.96 TeV
- **PYTHIA Perugia0**: soft QCD part is tuned using only minimum bias data from Tevatron and CERN ppbar data
- **PYTHIA DW**: uses the virtuality-ordered showers and used to describe the CDF II underlying events and Drell-Yan process data
- **PHOJET**: two-component Dual Parton Model with soft hadronic processes by Pomeron exchange and semi-hard processes by perturbative parton scattering
- **HERWIG+JIMMY**: cluster fragmentation model + MI interactions using JIMMY model
- **HERWIG++**: reimplemented in C++ cluster fragmentation model (+many new features)
- **PYTHIA ATLAS AMBT1**: P6 tuned by ATLAS to the low-multiplicity data

Main scope of comparison with Monte Carlo models: tune softQCD phenomenological models in order to use such models for better understood SM processes (pQCD, EWK measurements)





Towards a complete final state

- **Measurements using calorimeters have some advantages:**
 - Sensitive to a complete final state including neutrals (extra ~40%)
 - Independent systematics (compared to the tracking analysis)
 - Many high-precision jet measurements are based on energy deposition, and calorimeter-based UE studies can be directly used for such measurements
- **Calorimeter UE can take advantage of *unique ATLAS calorimeter*:**
 - 175k channels for electromagnetic, 5k channels for hadronic calorimeter
 - excellent transverse sampling and **longitudinal sampling(!)**

Excellent transverse and longitudinal segmentation allows reconstruction of “topological” clusters in 3D which are closely related to single particles (in average)

Topological clusters built from calorimeter cells

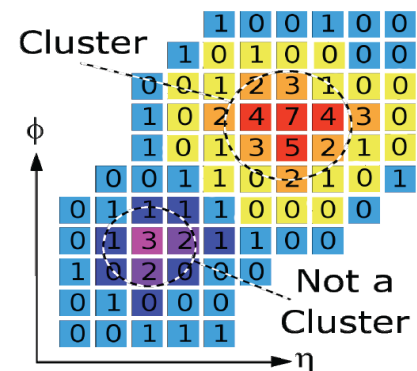
- follow shower development
- reduce noise and pile-up effects
- used for *jet reconstruction*

Method:

Seeded by cells with $|E| > 4 \times$ (noise level)

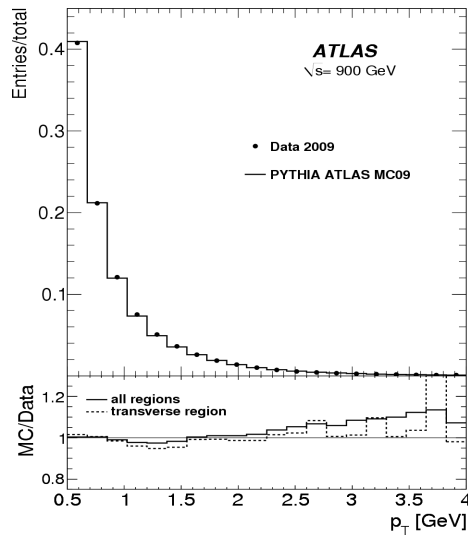
Neighboring cells with $|E| > 2 \times$ noise iteratively added (in 3D)

All neighbors around cluster ($|E| > 0$) added

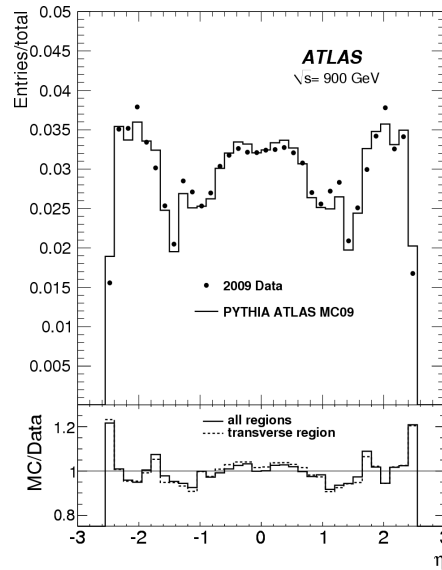




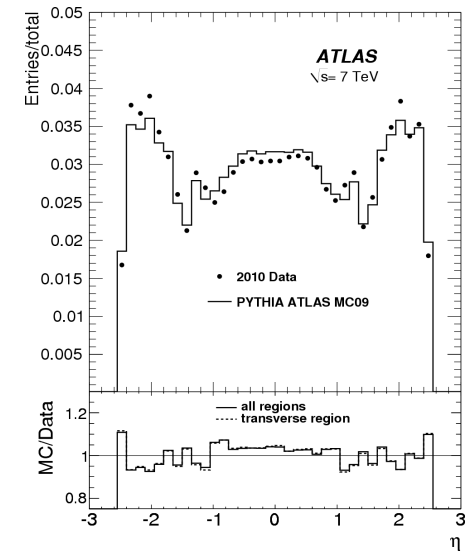
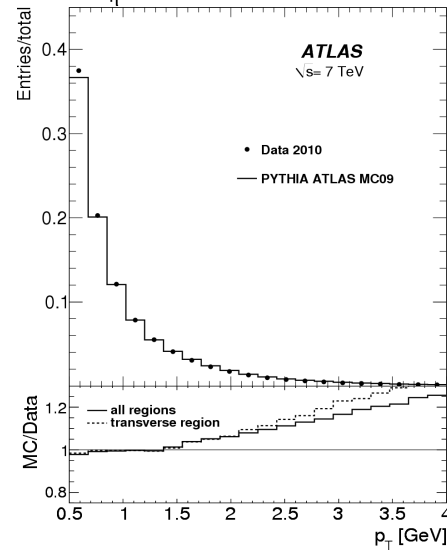
Topocluster properties



900 GeV



7 TeV



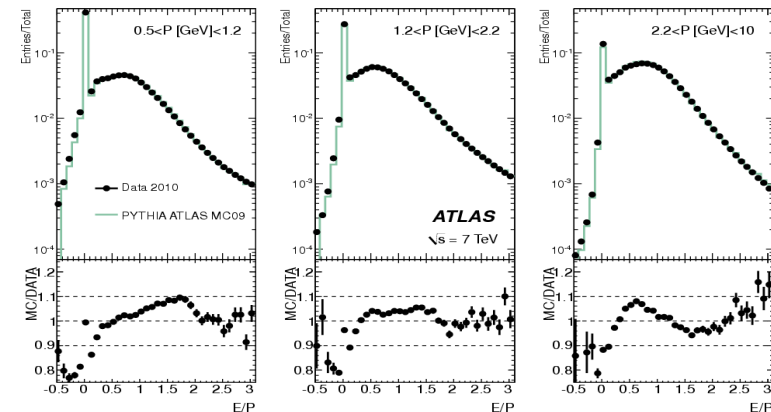
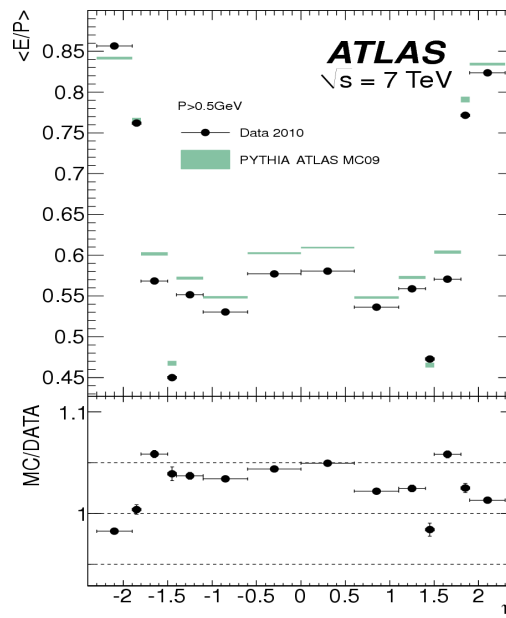
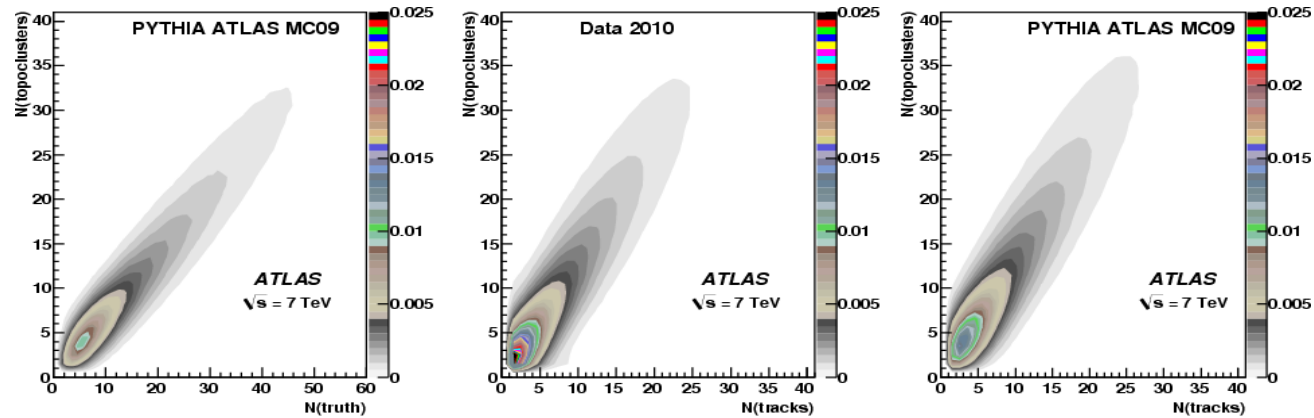
Plots from the paper draft





Measurements using topological clusters

- How well clusters correspond to individual particles? →
- Cluster overlaps
 - 2% effect for MinBias events
- How well Monte Carlo describe energies (relative energy scale) →
 - E/P
 - 5% uncertainty
- Simulation of clusters can be “calibrated” comparing ratios clusters/tracks for data and MC

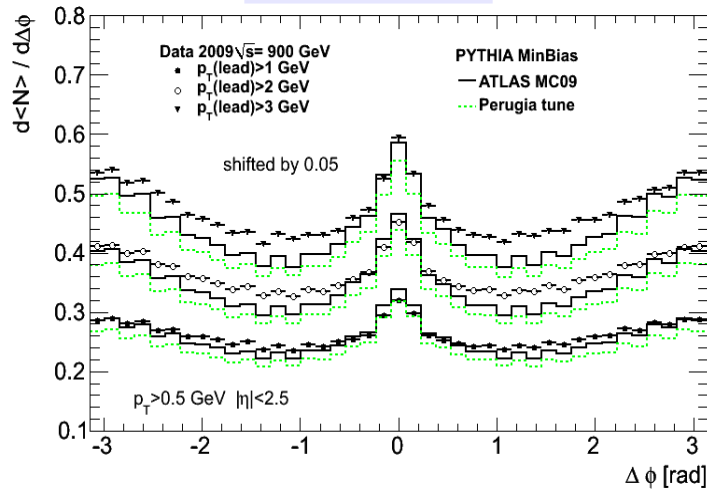


Unfold data to particle level using bin-by-bin corrections (~30-40%)

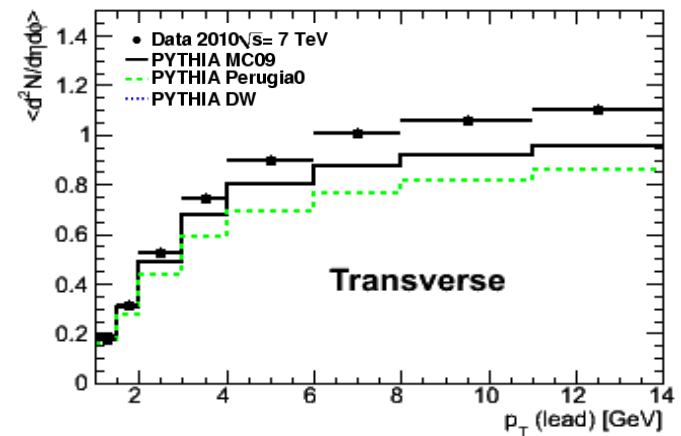
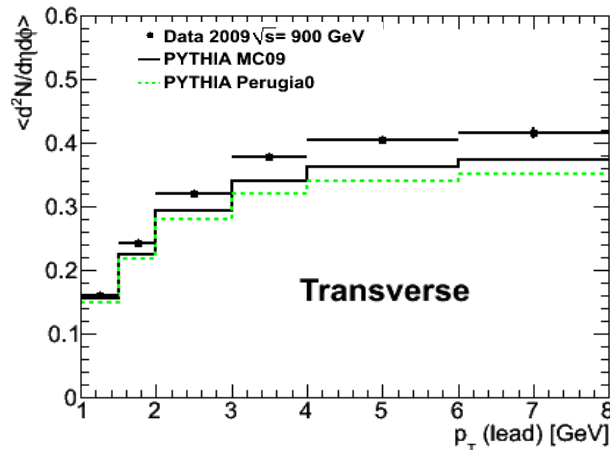
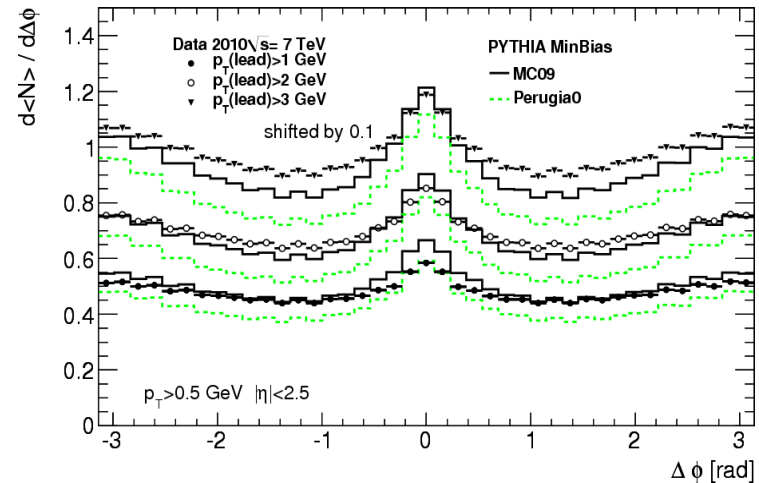
Plots from the paper draft

Detector-level distributions

900 GeV



7 TeV



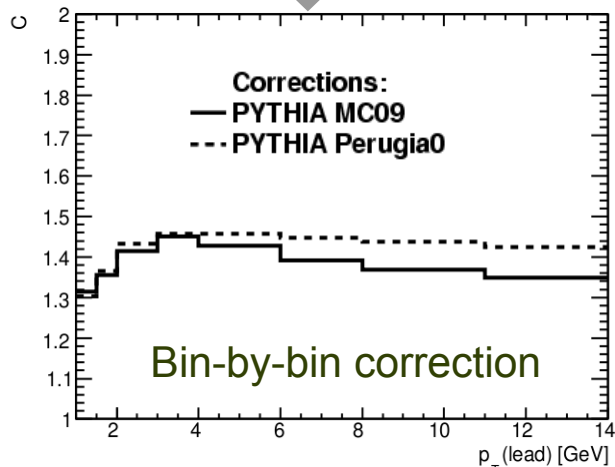
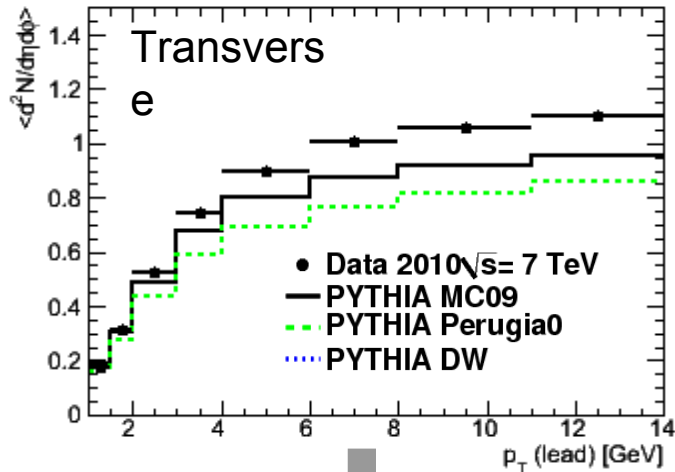
Sensitivity to MC tunes. Can be used for MC tuning

- unfold the distributions to the truth level to simplify the task

No single MC tune with a good description for all distributions.

Example of the correction procedure

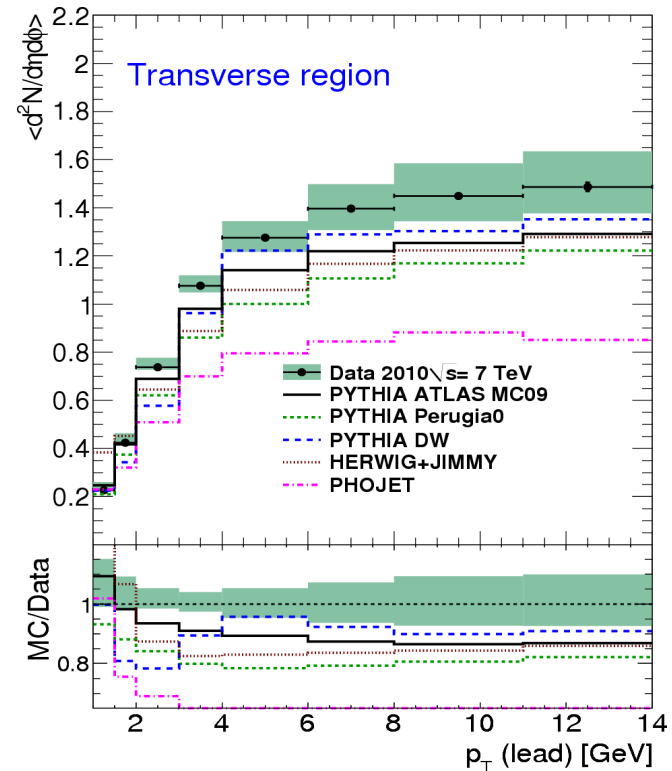
detector-level



(Perugia0 and MC09 have different $p_T(\text{lead})$ spectra)



Particle level (+ systematics)



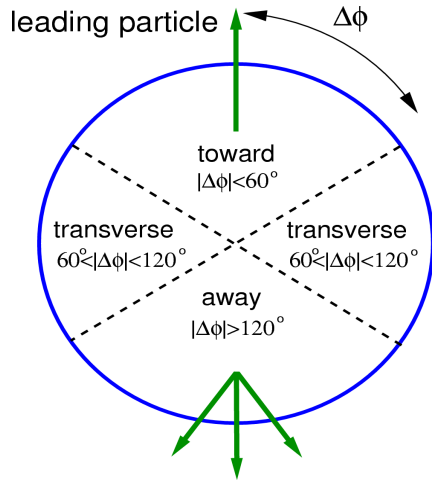
Bin-by-bin correction:

$$C = N(\text{truth})/N(\text{reco})$$

- Validated using track-based UE/MinBias studies
 - ATL-COM-PHYS-2010-165 and ATL-COM-PHYS-2010-237

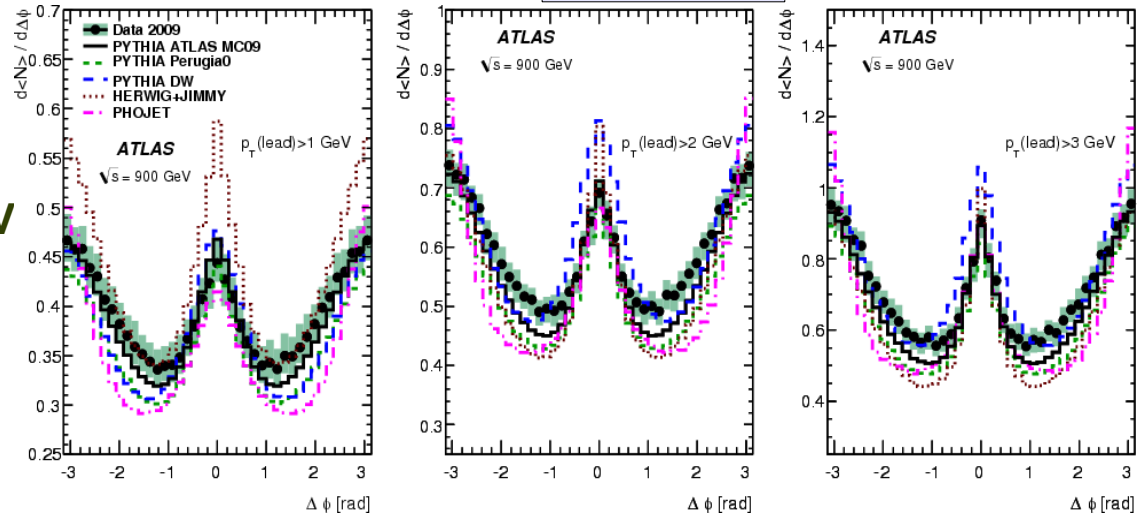


Densities as a function of $\Delta\phi$ ($p_T > 100$ GeV)

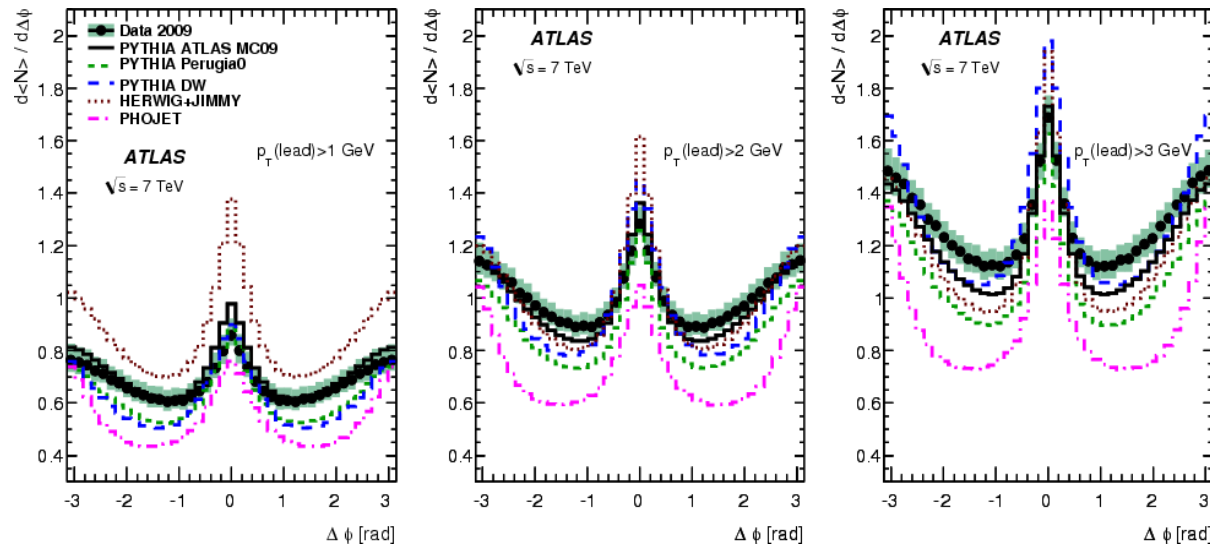


$p_T > 500$ MeV

900 GeV



7 TeV



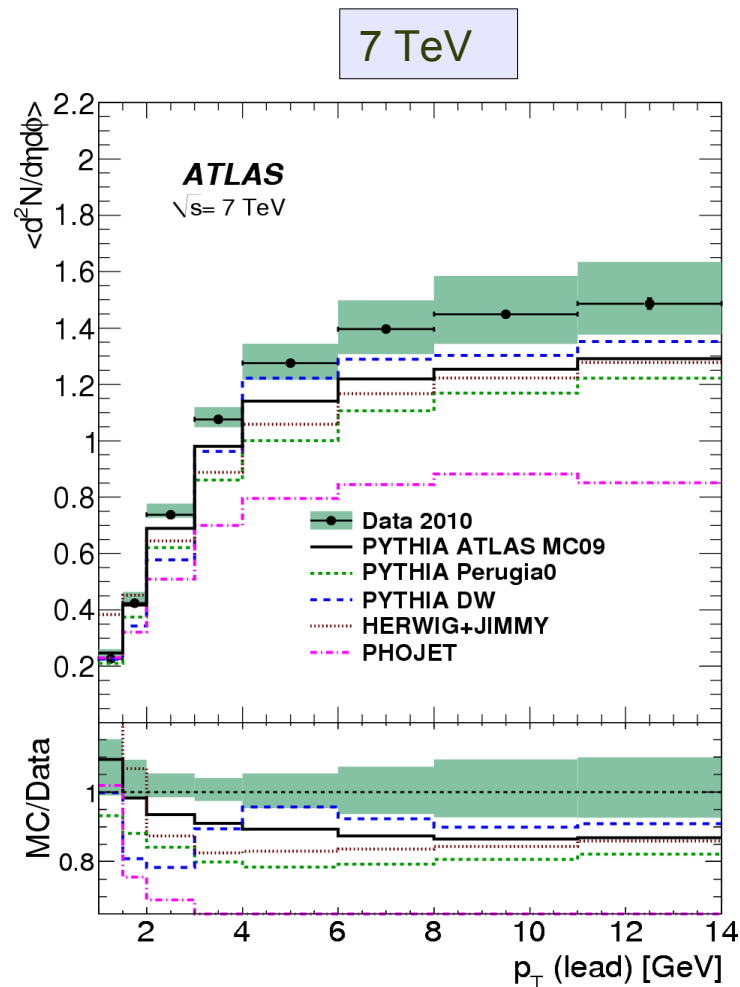
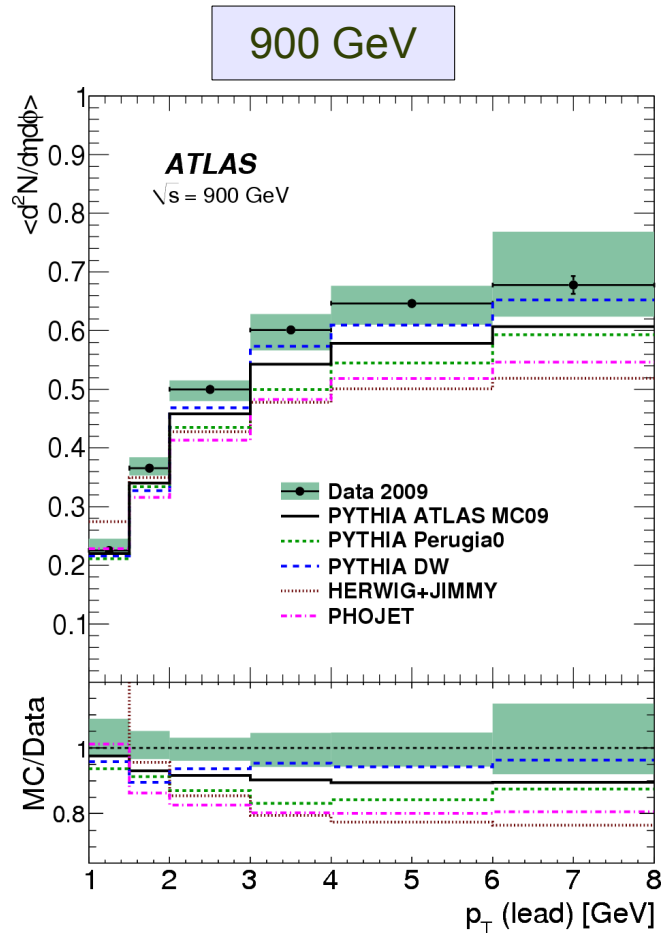
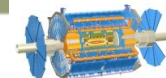
All Monte Carlo models fails
(similar to tracking analysis)

~40% increase in particle
density due to neutrals

Plots from the paper draft



Particle densities in the transverse region

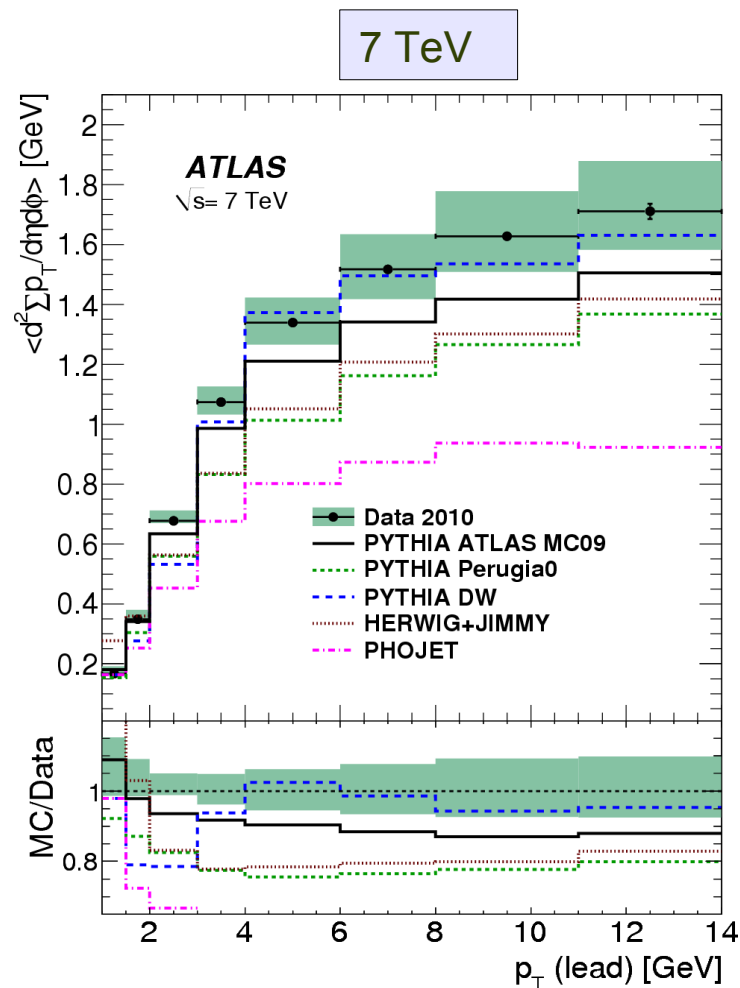
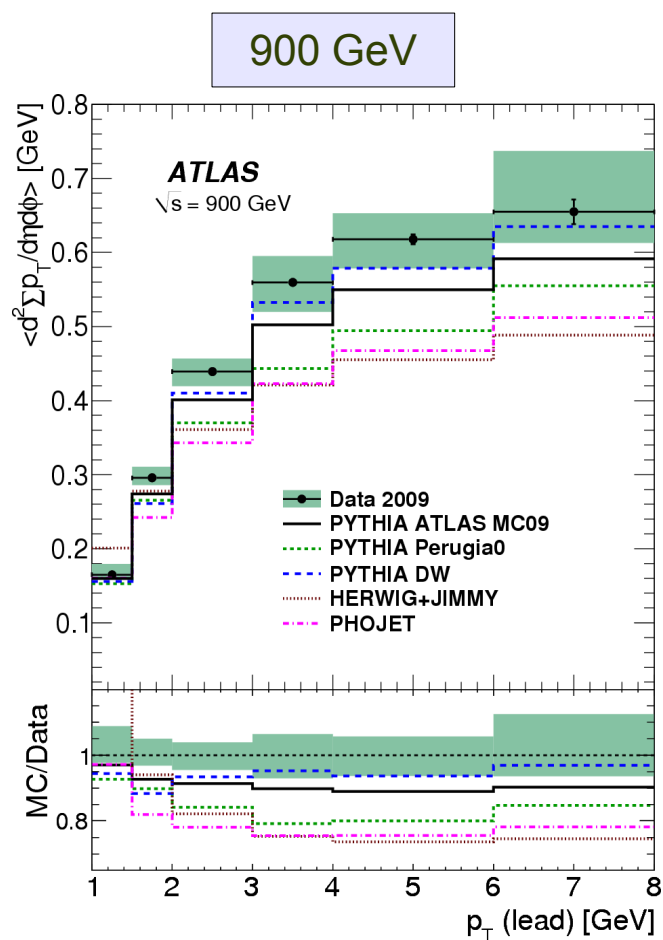
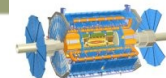


Lower particle densities in Monte Carlo simulations
Larger discrepancies at 7 TeV (PHOJET fails completely)

Plots from the paper draft



Scalar pT sum in the transverse region



Lower particle sum pT densities in Monte Carlo simulations
Larger discrepancies at 7 TeV (PHOJET fails completely)

Plots from the paper draft



Systematic uncertainties

- Reject low-multiplicity events
- Energy scale using the grid in η -P (to take into account 10% uncertainty in the transition region)
 - Includes +3 MeV shift to account for the difference between data and MC for π^0 peak
- ± 0.025 rad for cluster centers ϕ and η (shift by 1 Ecell)
- (a) PYTHIA with 10% extra material; (b) with improved PP0 geometry
 - <http://cdsweb.cern.ch/record/1243587>
- Using Perugia0 for unfolding (model dependence, difference in $p_T(\text{lead})$ spectra)
- Multiplicity of clusters vs MC tracks was re-weighted
- Resolution tails were re-weighted
- Entire analysis repeated using EM-scale clusters
- Alternative hadronic shower in GEANT (FTFP-Bertini)
- Removing gap region

Table shows max uncertainties (not average)

Check	$d \langle N \rangle / d\Delta\phi$	$\langle d^2 N / d\eta d\phi \rangle$	$\langle d^2 \Sigma p_T / d\eta d\phi \rangle$
$N < 3$ rejection	+0.2%	+10%	+10%
Energy scale	$\pm 4.3\%$	$\pm 4\%$	$\pm 5.6\%$
ϕ position	$\pm 1.3\%$	$\pm 0.2\%$	$\pm 0.2\%$
η position	$\pm 0.2\%$	$\pm 0.2\%$	$\pm 0.2\%$
Additional material	+3.5%	+3%	+3.6%
Model dependence	$\pm 3.5\%$	$\pm 5\%$	$\pm 4.5\%$
Multiplicity reweighting	$\pm 4.5\%$	$\pm 10\%$	$\pm 11\%$
Resolution reweighting	$\pm 0.4\%$	$\pm 6\%$	$\pm 6\%$

Detailed budget of uncertainties in each bin for each observable is given in the notes

Summary

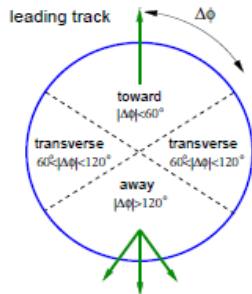
- **First UE measurements using calorimeter clusters**
 - Important measurement for jet-related studies where UE should be understood
 - Studies are sensitive to the entire hadronic final state:
 - densities are ~40-45% larger compared to charged-particle UE measurements
- **Can enable studies of jet-shapes and jet-substructure for boosted jets where understanding of topoclusters inside jets and unfolding to particle level are of primary importance**
- **Provide systematically independent check of track-based UE measurements**
 - Additional constraint on the understanding of UE & model tunings
- **TopoClusters measurements confirm the conclusions for charged-particle UE studies for pre-LHC MC tunes**
 - No MC tunes with good description of all observables
 - MC models have smaller particle activity in the transverse region



Backup



Towards a complete final state



Goal: to reconstruct densities in the transverse region which are

- Based on a complete final state (extra 40%)
- Closer related to future jet measurements
- Systematically independent of the tracking analysis

Use calorimeter clusters as objects which are most directly related to separate particles

The central question is not how to associate particles to clusters:

- studied over the past 5 years!
- but what systematical uncertainties should be attributed to such association for a concrete physics measurement?
- How to treat overlaps?

All such questions are vital for many future physics topics
(event shapes, jet shapes for boosted jets etc. etc.)

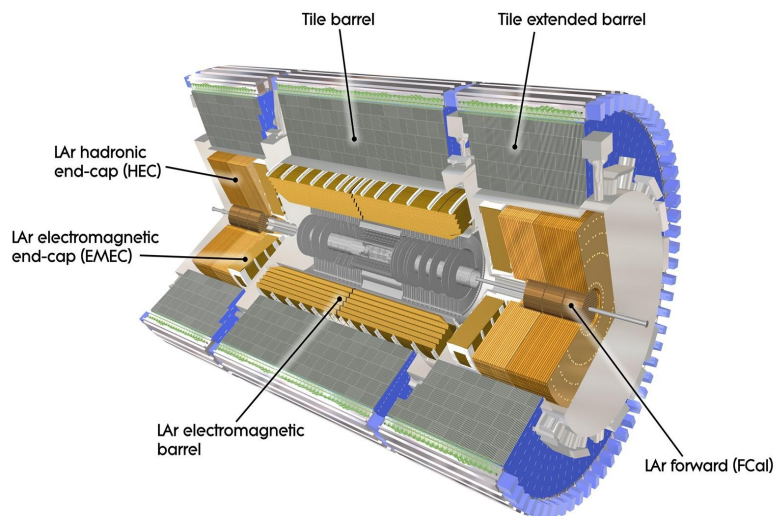
Example:

MinBias events have very small cluster overlaps:

- sparse events, with ~ 10 clusters per event
- probability of having 2 particles in the cone 0.2 in η - ϕ $< 1\%$
- looking at average multiplicities

2 particles \rightarrow 1 cluster

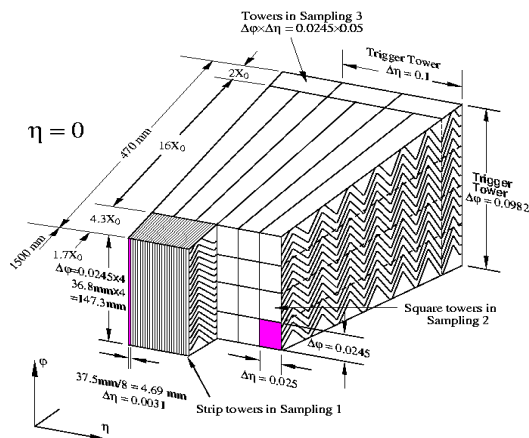
Fully take advantage of unique ATLAS calorimeter



- ~ 175k channels for EM
- ~ 5k channels for hadronic calorimeter
- excellent transverse sampling
- **longitudinal sampling**

No other experiment @ LHC has longitudinally segmented calorimeter

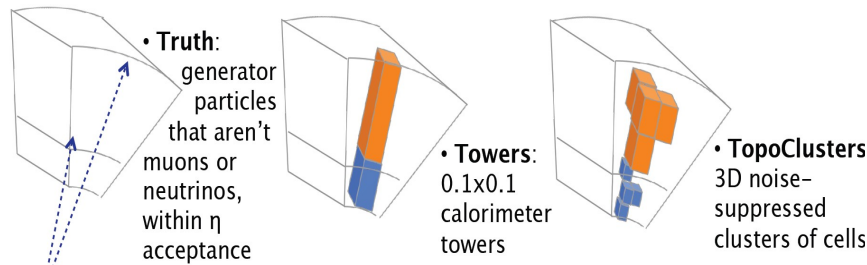
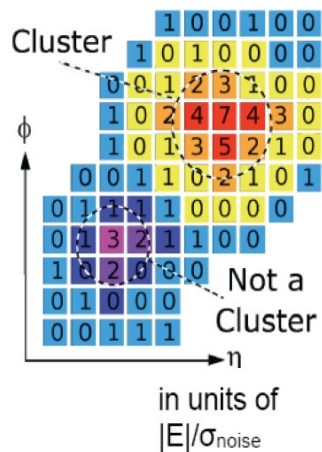
- CMS calorimeter has no longitudinal segmentation, less channels, smaller thickness



Excellent transverse and longitudinal segmentation allows reconstruction of “topological” clusters in 3D which are closely related to single particles (in average)

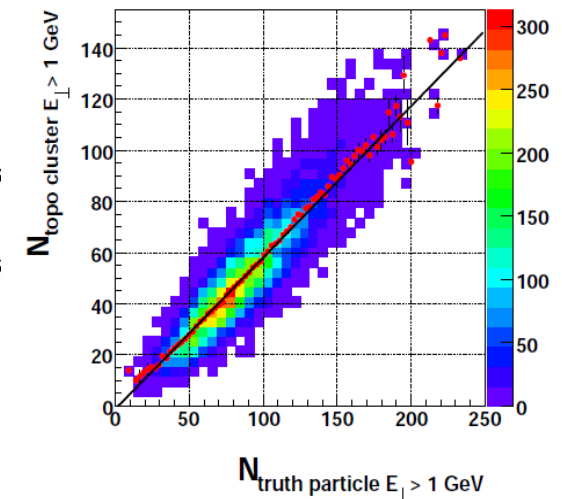
UE studies using topoclusters

- Use calorimeter measurements taking advantage a fine calorimeter granularity
- Topoclusters are the natural choice for such measurements:
 - provide efficient noise and pile-up suppression
 - follow shower development → correspond to individual hadrons
 - collect nearest neighbors around seed cells with a signal above 4σ above noise
 - neighboring cells collected to a cluster if the signal significance is above a secondary seed threshold 2σ
 - collect all surrounding cells if no further secondary seeds are found



(From a P.Loch's talk)

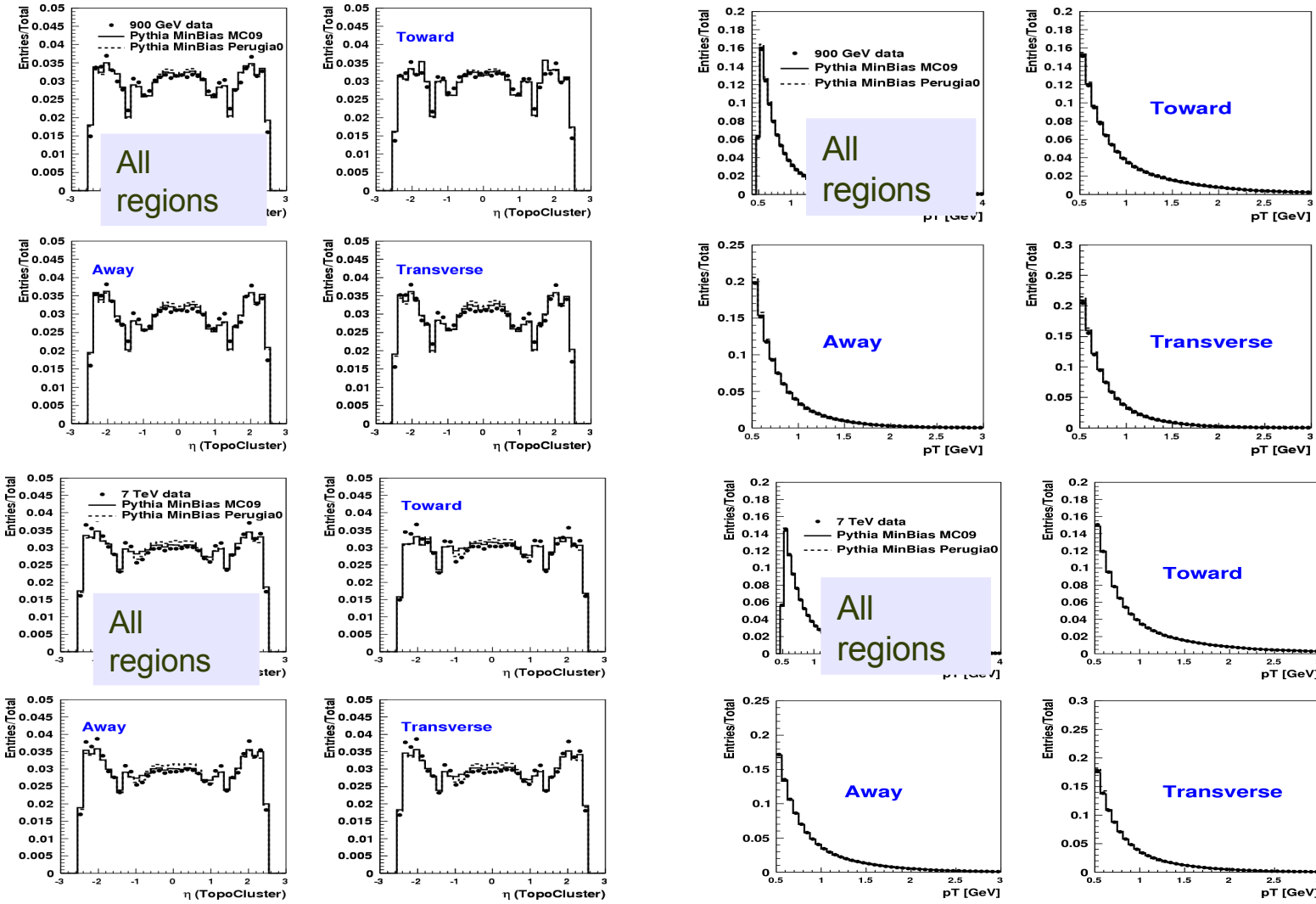
1 cluster corresponds to 1.6 truth particles



(S.Menke talk, 2008)

- The UE measurement concentrates on
- “density” measurement rather than “energy measurement”
 - somewhat reduced energy scale uncertainty

TopoCluster properties at 900 GeV and 7 TeV



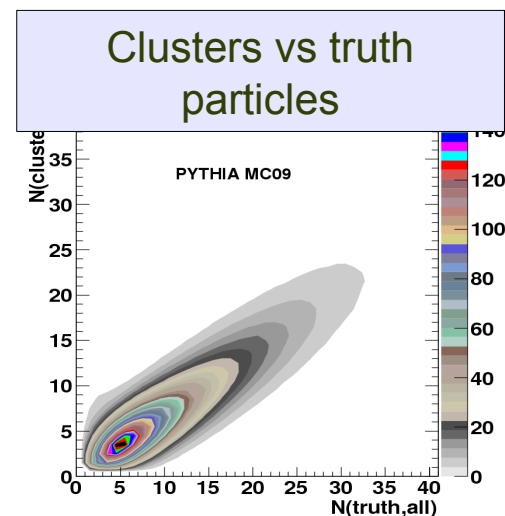
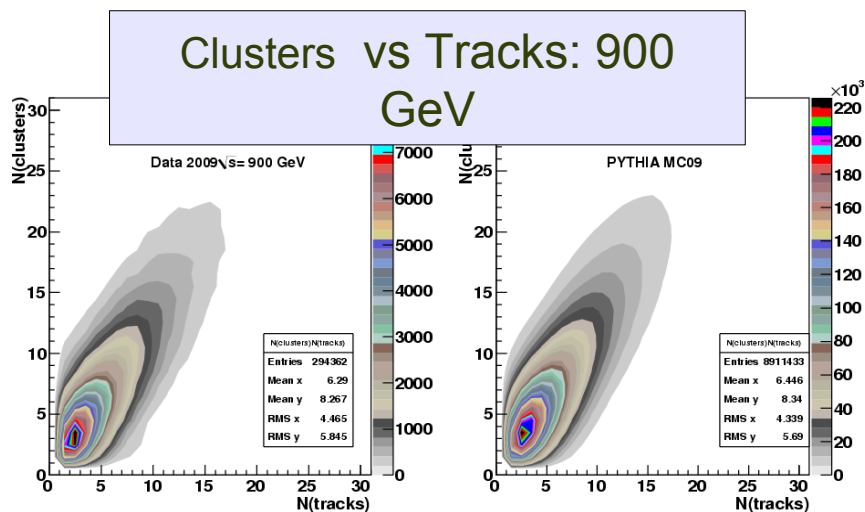
900
GeV

7 TeV

- Reasonable agreement with MC09 & Perugia0 tunes

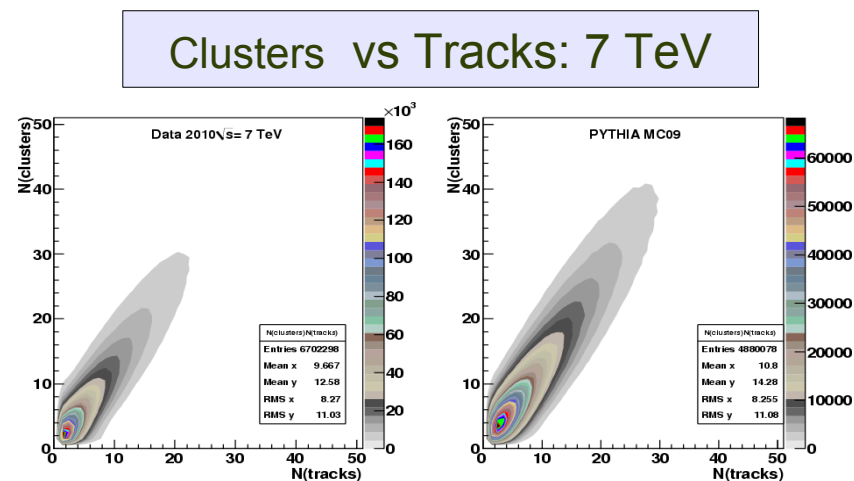
Topocluster properties for the UE studies

- Good association with the number of truth hadrons
- Reasonable description of data using MC



Although visually identical, and Kolmogorov–Smirnov gives $p \sim 0.999$, data and MC still have discrepancies

Small discrepancies ($\sim 10\%$) is taken into account by re-weighting MC09

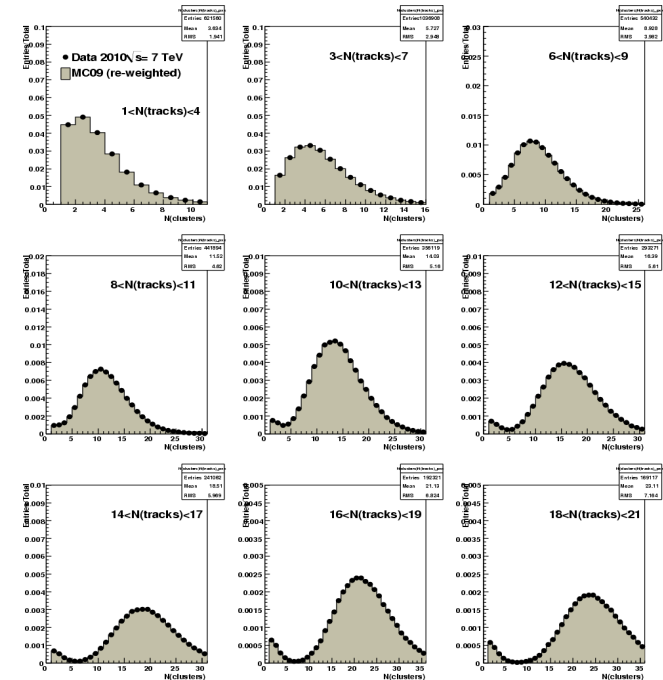


Topocluster performance

To correct topocluster observables to particle level we must be sure that MC reproduces every aspect of topoclusters properties (at least, in average)

This is done by comparing topoclusters with tracks (for charged-particle case) or using some reference measurements for neutral topoclusters (like p_0).

- **Comparing MC multiplicities of topoclusters with data (for fixed number of tracks)**
 - controls many performance aspects, including possible threshold effects, overlaps, etc.
 - MC09 was re-weighted in bins $N(\text{track})$ to match the data
- **Comparing data and MC for $\langle E/p \rangle$**
 - relative energy scale uncertainty (enters via $p_T(\text{min})$ cut)
 - propagated to the final measurements
- **Comparing resolution distributions for leading clusters (i.e. $p_T(\text{clusters})/p_T(\text{tracks})$)**
 - MC09 re-weighted to match the data
- **Plus 10 cut variations to understand various aspects of topocluster performance, alternative MC, material map, shower models etc.**



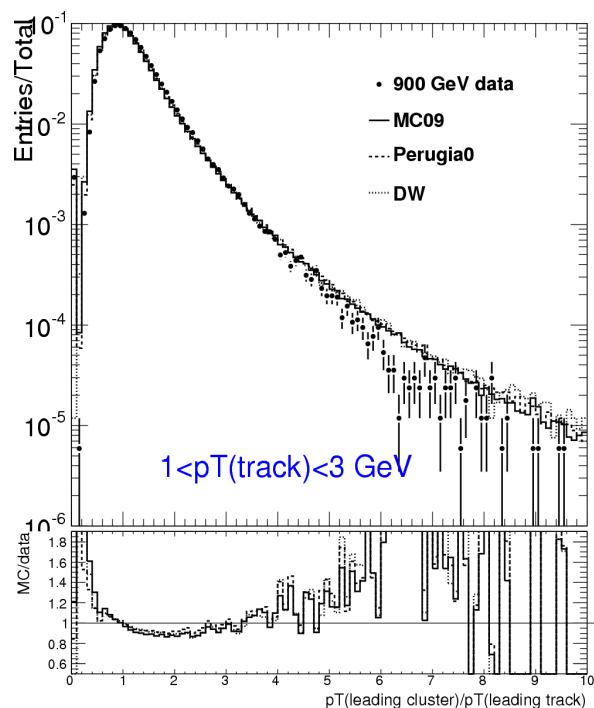
Example: distributions of topoclusters for events with a fixed number of tracks used for re-weighting

Resolution tails for leading topoclusters

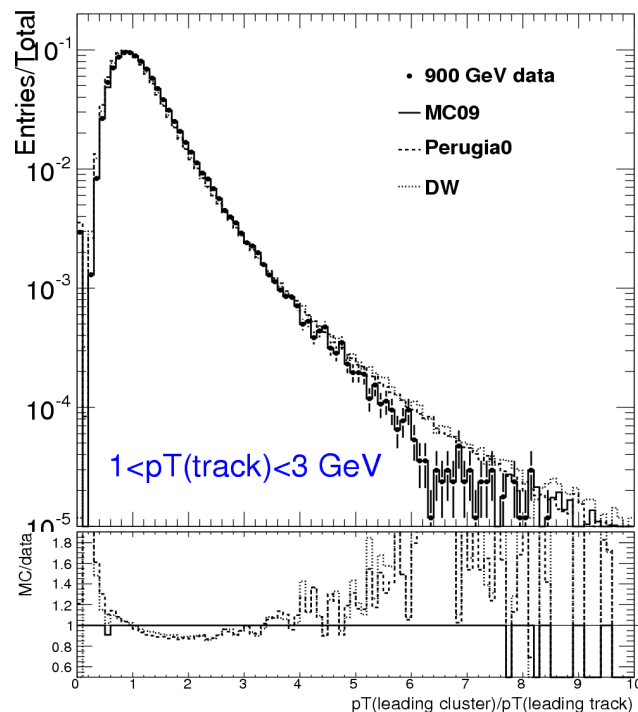
Migration effects from low-pT to high-pT should be well described by a MC

One way to control it is to look at the tail $pT(\text{cluster,lead})/pT(\text{track,lead})$ and re-weight MC

Example:



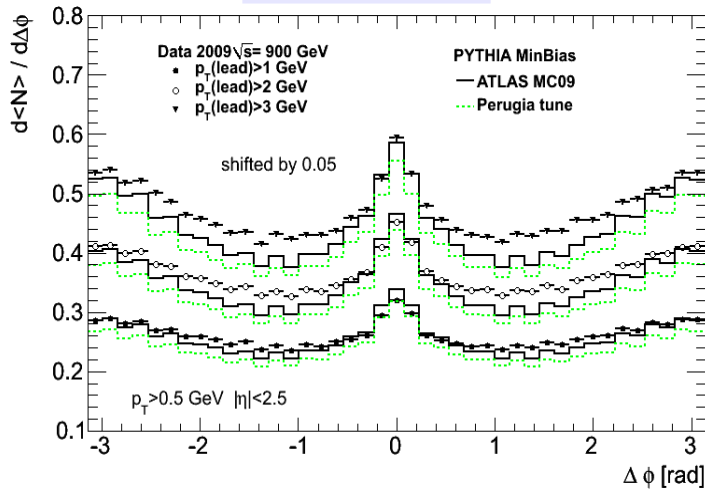
MC09
re-weighted



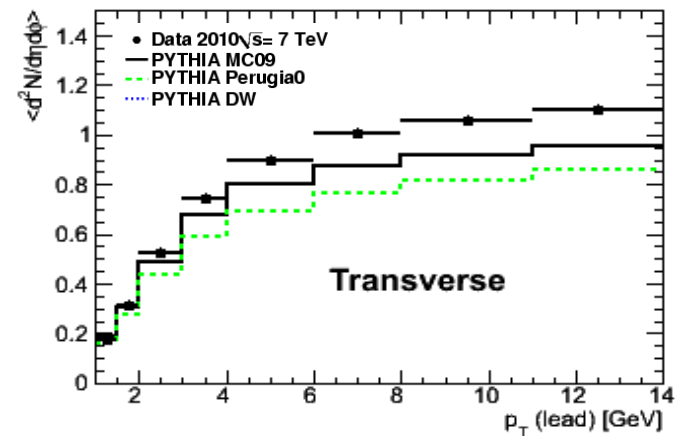
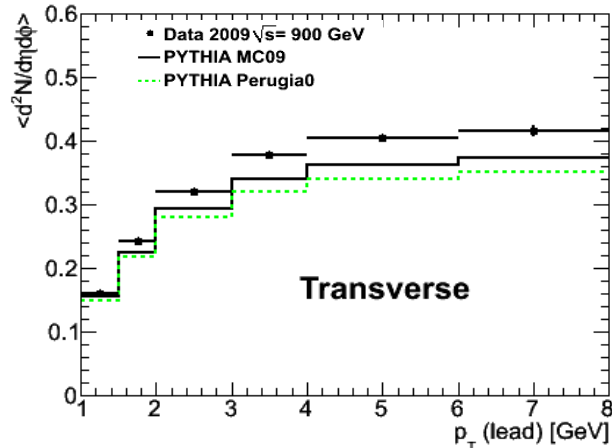
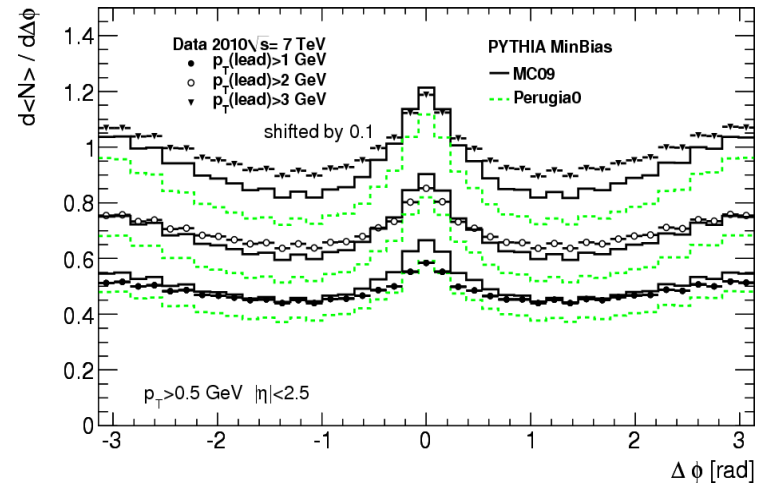
re-weighting is done in bins of pT for 900 GeV and 7 TeV data separately

Detector-level distributions

900 GeV



7 TeV



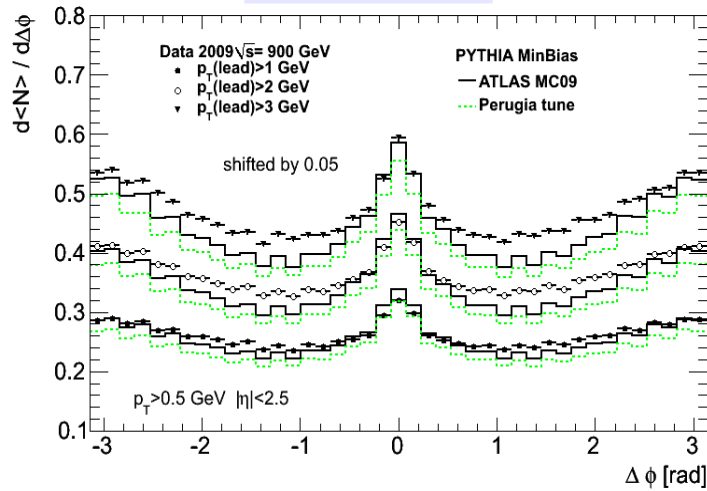
Sensitivity to MC tunes. Can be used for MC tuning

- unfold the distributions to the truth level to simplify the task

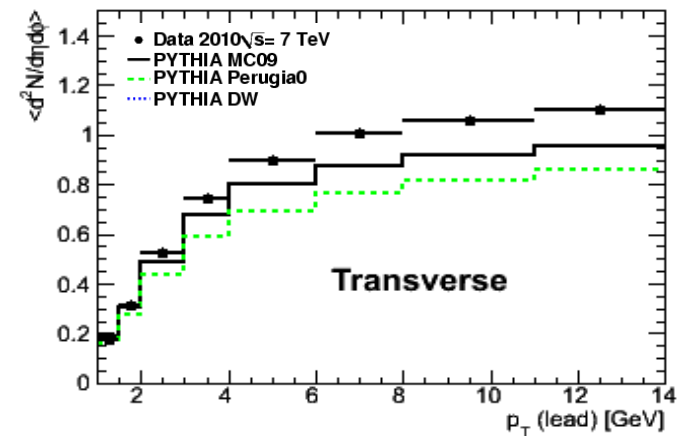
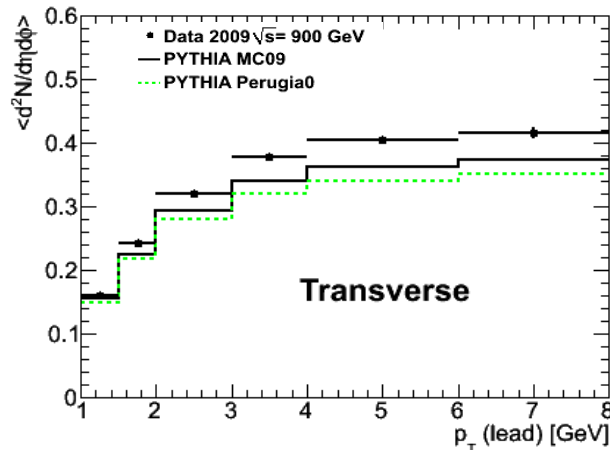
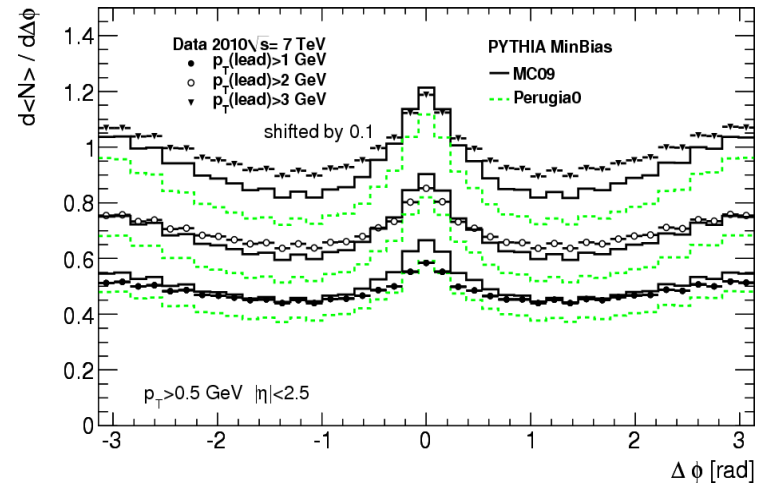
No single MC tune with a good description for all distributions.

Detector-level distributions

900 GeV



7 TeV



Sensitivity to MC tunes. Can be used for MC tuning

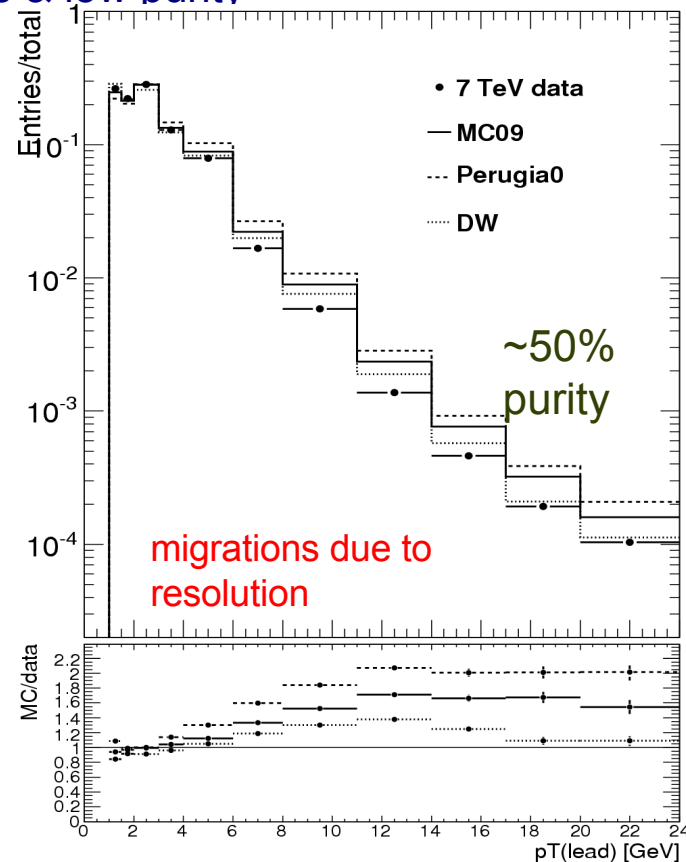
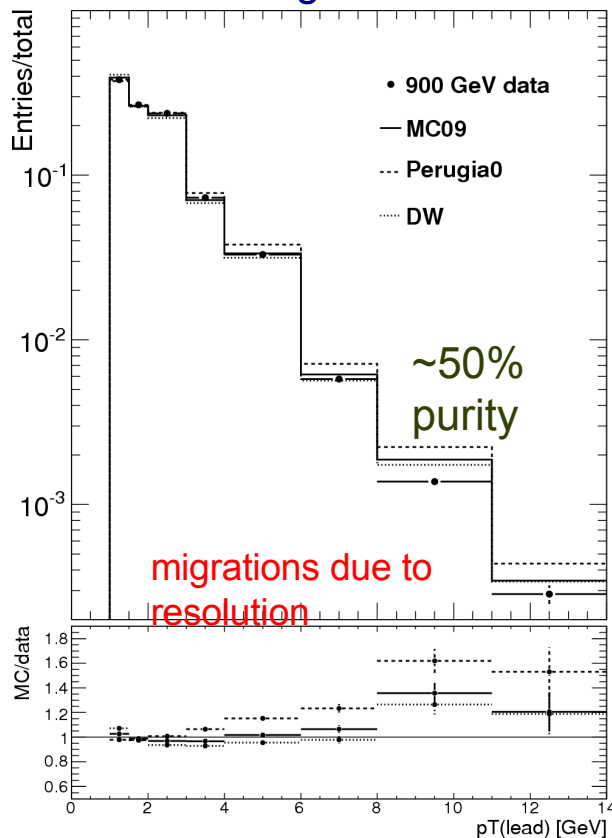
- unfold the distributions to the truth level to simplify the task

No single MC tune with a good description for all distributions.

Resolution tails

- Must be sure that resolution tails for leading clusters are well described
 - Previous slides show that the resolution is well described for all clusters
 - Since we take leading clusters at large p_T we probe the detector performance in greater details (see the Pisa workshop material, “Using leading topoclusters to probe detector performance”)
 - Resolution tails drop slower with p_T than jet cross section
 - Leads to large rate of fake topoclusters & low purity

Difference between
different MC tunes



Correction procedure

- All distributions are measured with respect to “reference” particles
- Mismeasured particle introduces smearing (lower purity in bins)
- Can be taken into account using a bin-by-bin correction:

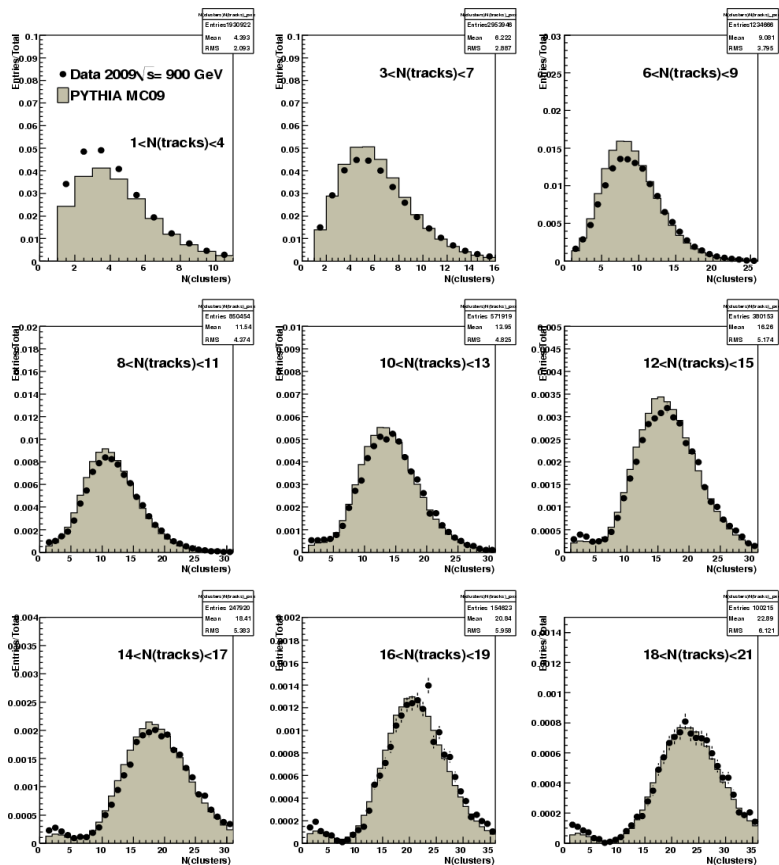
$$C = N(\text{gen})/N(\text{reco}) = \text{purity} / \text{efficiency}$$

Corrects for:

- event selection
- clusters selection (inefficiencies due to threshold cut effect, losses, merging/splitting etc.)
- resolution smearing (leading cluster is lost), other impurity effects
- decays of long-lived resonances (truth level is defined by $\tau < 3 \cdot 10^{-10}$ sec)
- Resolution smearing is minimized choosing bin sizes larger than resolutions in each bin
- Model dependence is controlled using alternative MCs
- Tested using track-based MinBias studies (fully agrees with the track-weighting approach)
 - ATL-COM-PHYS-2010-165 and ATL-COM-PHYS-2010-237
- No correction for diffraction was applied:
 - Single and Double diffraction is expected at the level of:
 - <1% for PYTHIA (SD/DD) when $p_T > 1$ GeV
 - ~1% for PHOJET (SD/DD) – more diffractive events at $p_T > 1$ GeV (hard diffraction), but SD/DD are similar in shape and show a small contribution to the final densities
- Only measurements are presented where the correction factor are understood and <50%

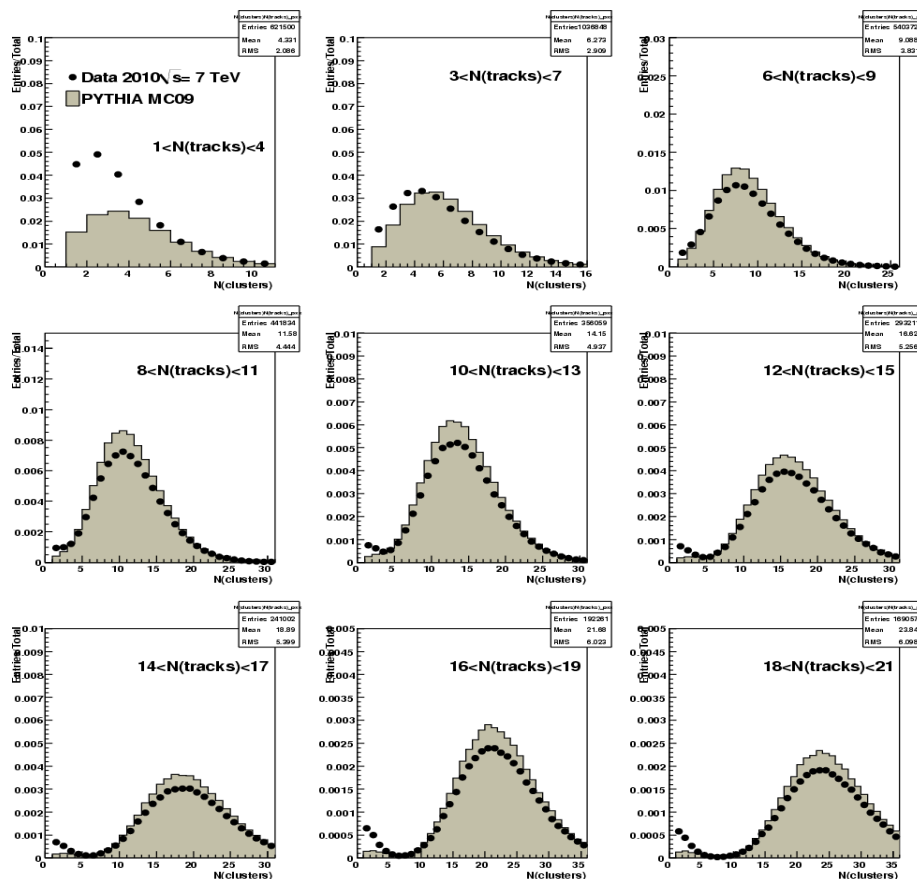
Topoclusters vs Tracks. Slices after 2D normalization

900 GeV



- Discrepancies for low multiplicities
- Likely reason - diffraction.

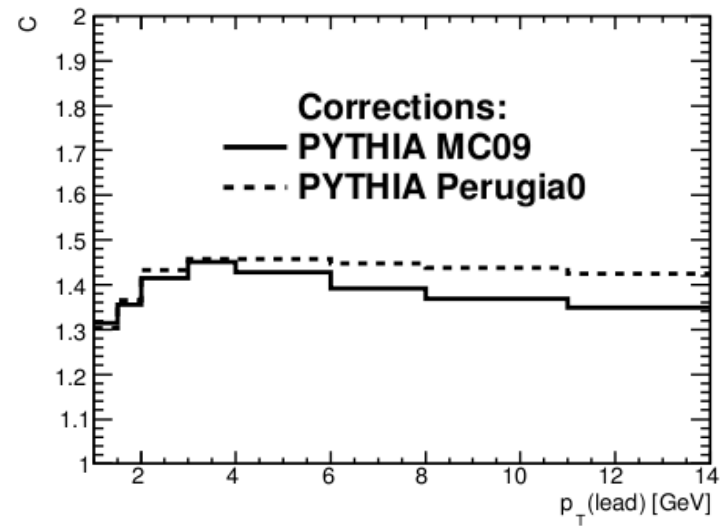
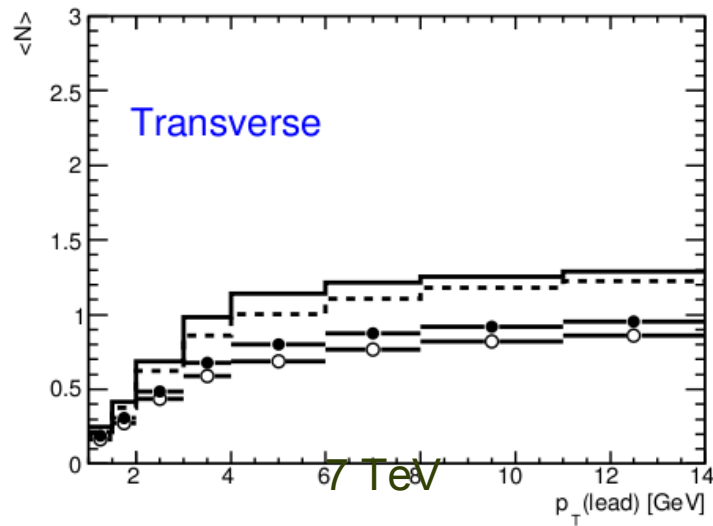
7 TeV



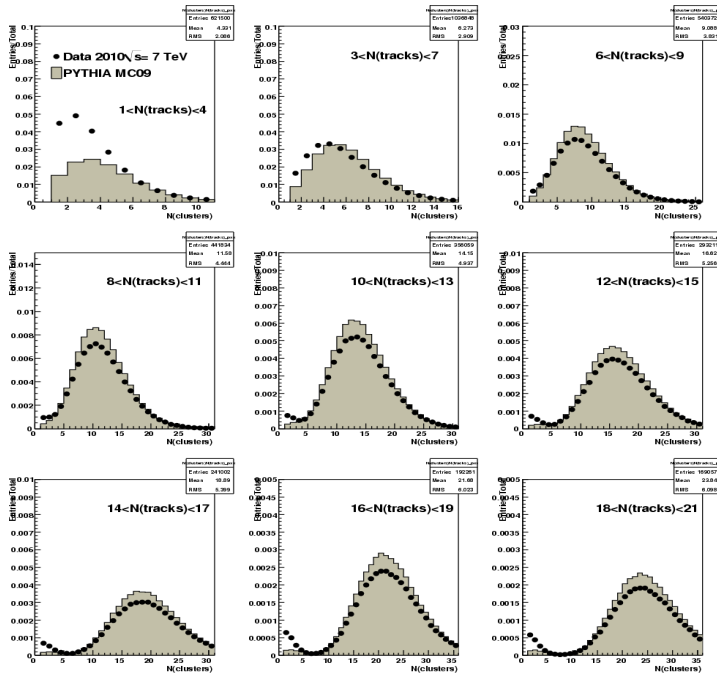
- discrepancies will be taken into account reweighing events, i.e. using data/MC as weights

MC dependence

- Clusters, MC09
- Truth, MC09
- Clusters, Perugia0
- - - Truth, Perugia0

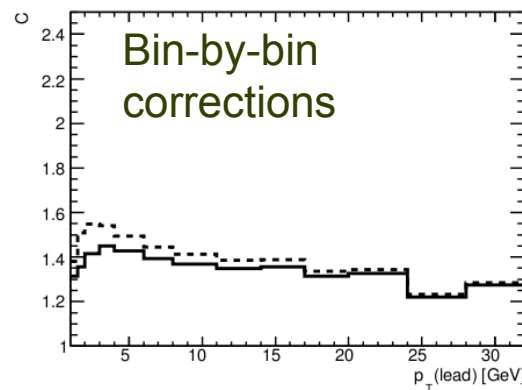
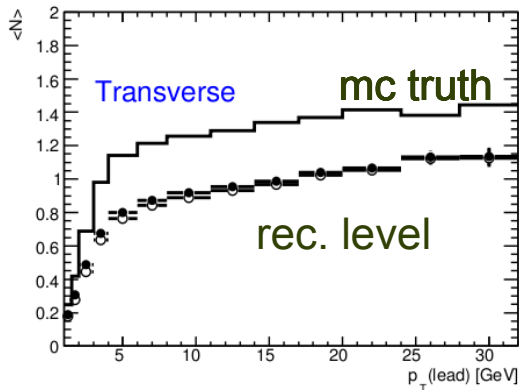
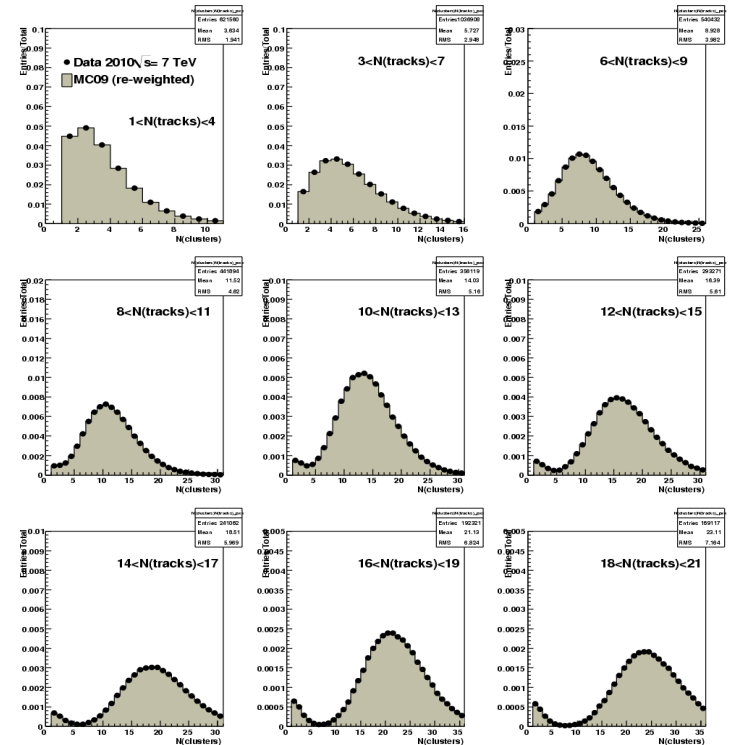


Multiplicity re-weighting. 7 TeV



re-weight

Data/MC

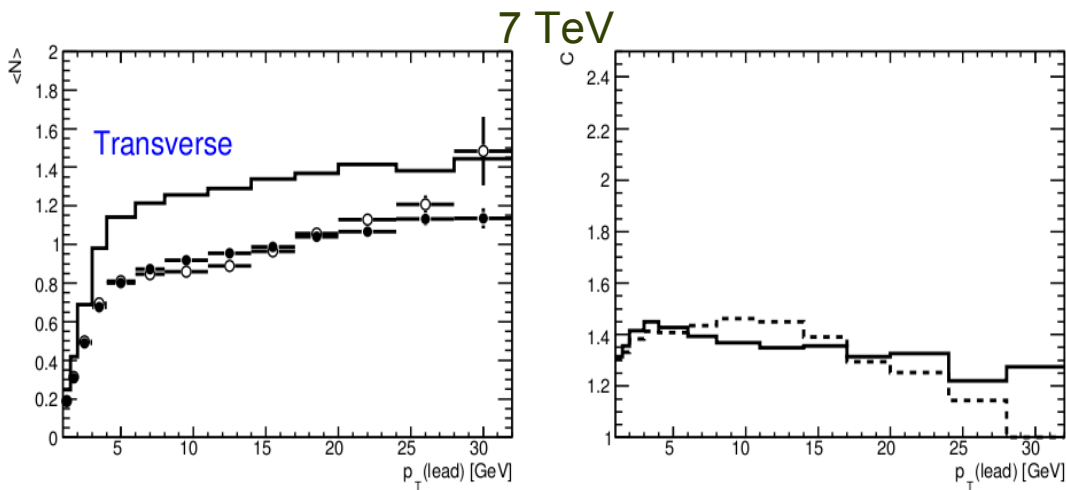
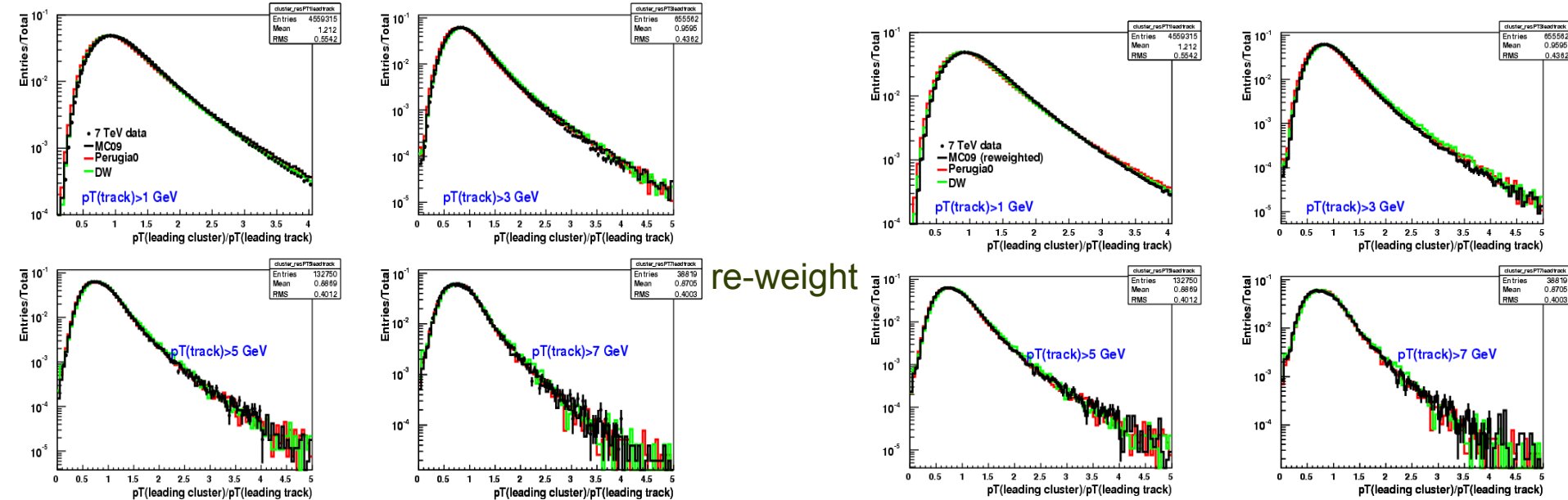


Solid line: before re-weighting

Dashed: after re-weighting

Note: effect from re-weighting for 900 GeV is smaller

Resolution tails for leading clusters re-weighting



Solid line: before re-weighting

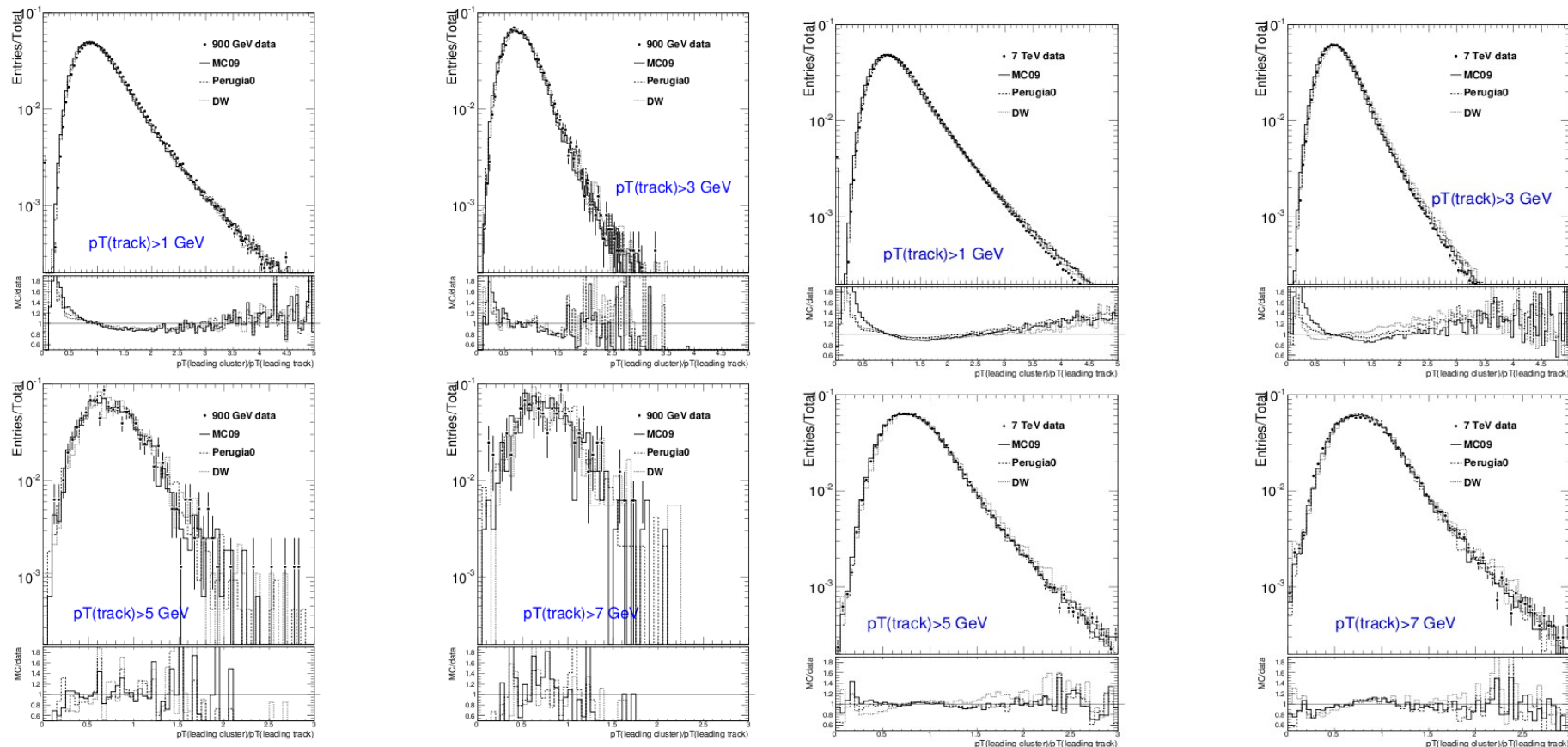
Dashed: after re-weighting

Resolution tails for leading topoclusters

For a given leading track, identify a leading cluster and look at $pT(\text{cluster,lead})/pT(\text{track,lead})$
Should be relatively independent of MC tune. How well we understand the resolution tails?

900 GeV

7 TeV

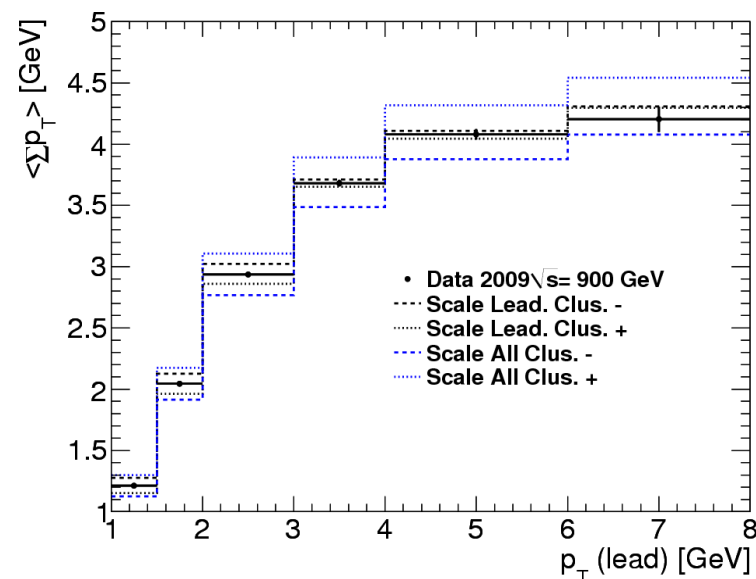
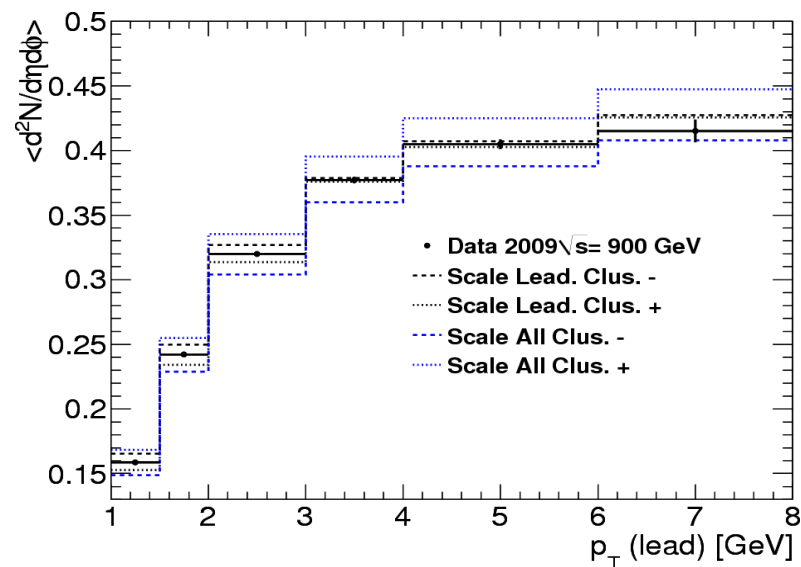
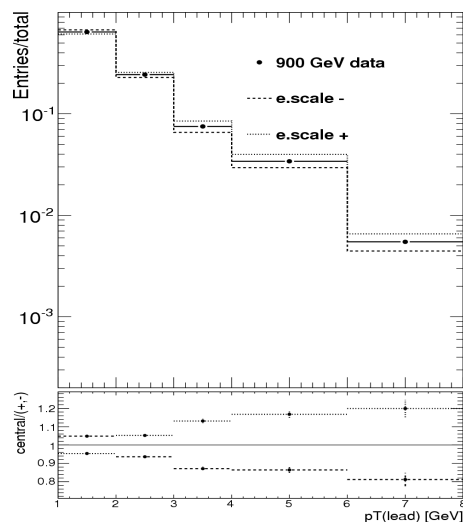
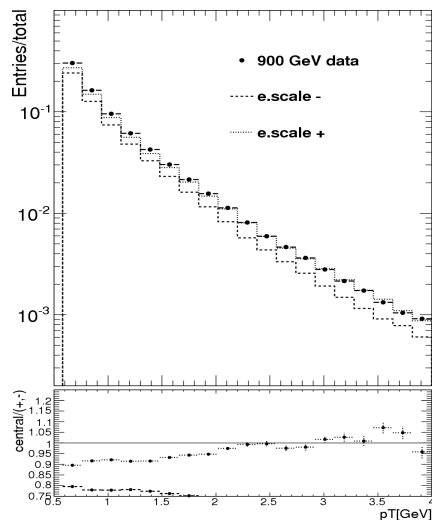


Discrepancies between data and MC should be taken into account

Due to some MC-tune dependence, correction to the resolution tails will be used for systematics (rather than for the central values)

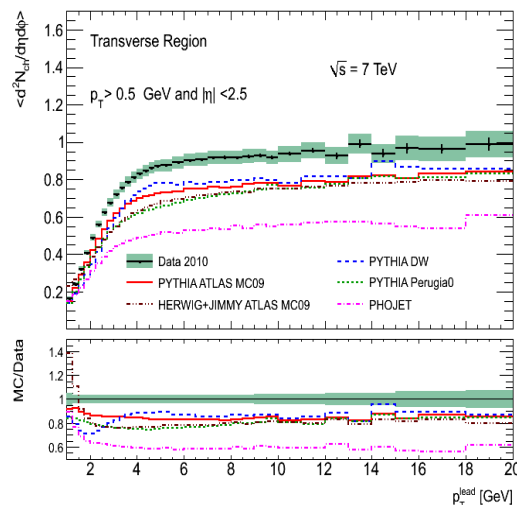
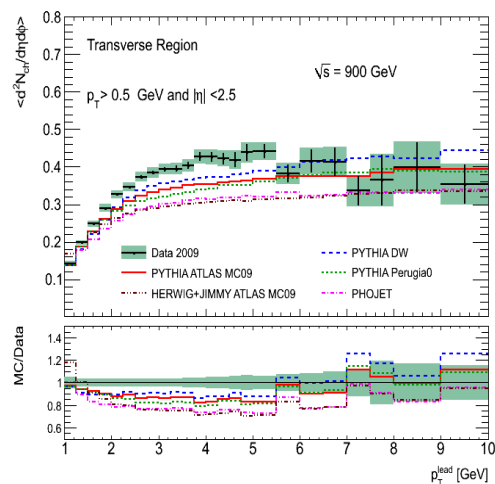


Energy scale uncertainties

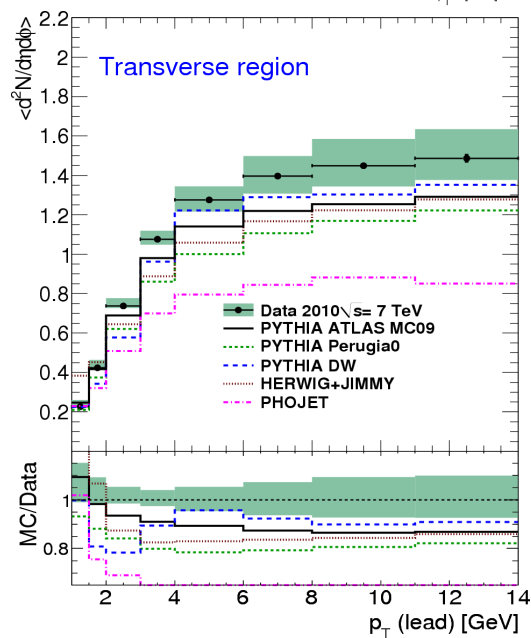
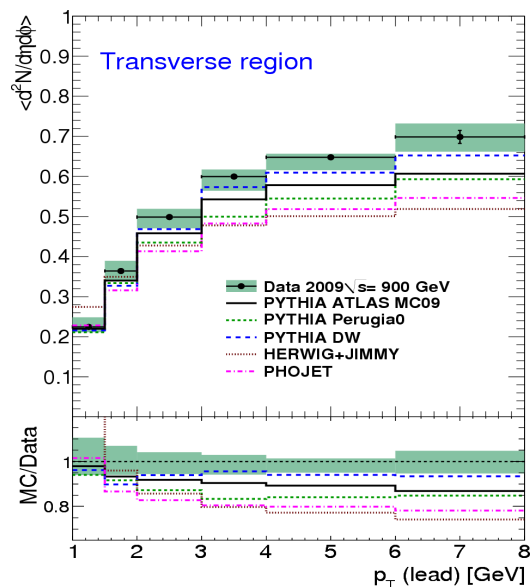


Partial cancellation of scale uncertainties for $\langle N \rangle$ vs $p_T(\text{lead})$

Comparing with the tracking results



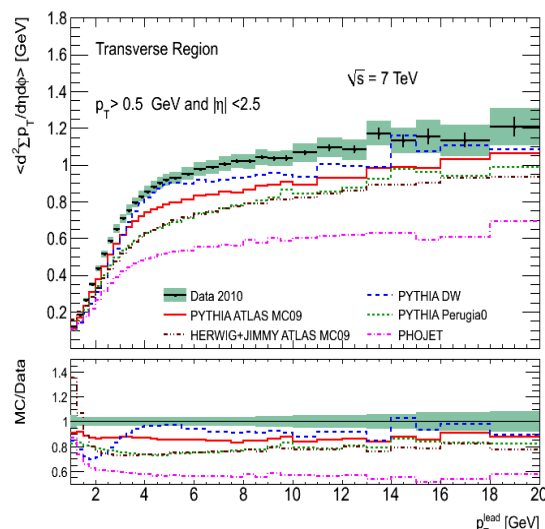
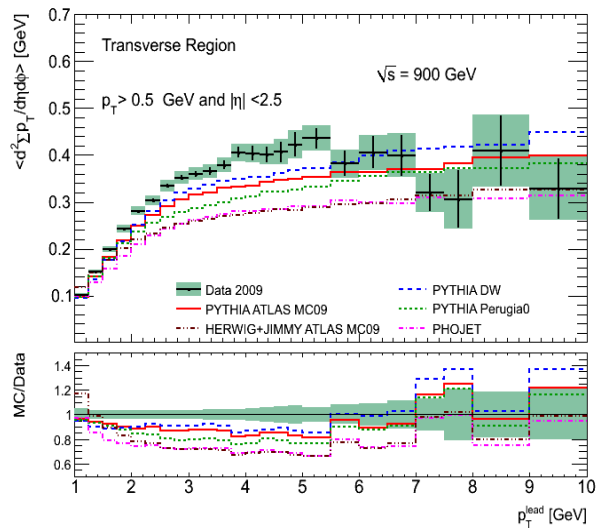
Charged particles



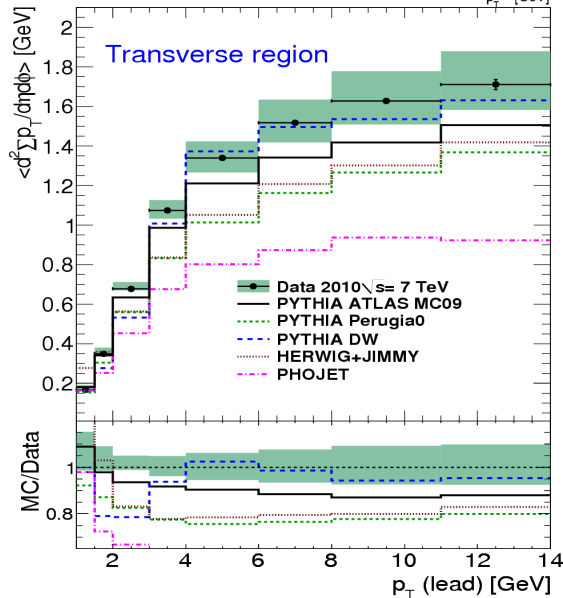
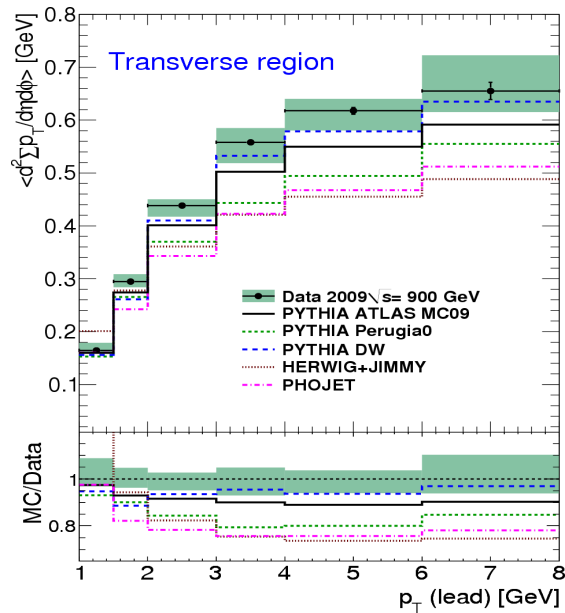
All particles

densities increased by ~40%

Comparing with the tracking results



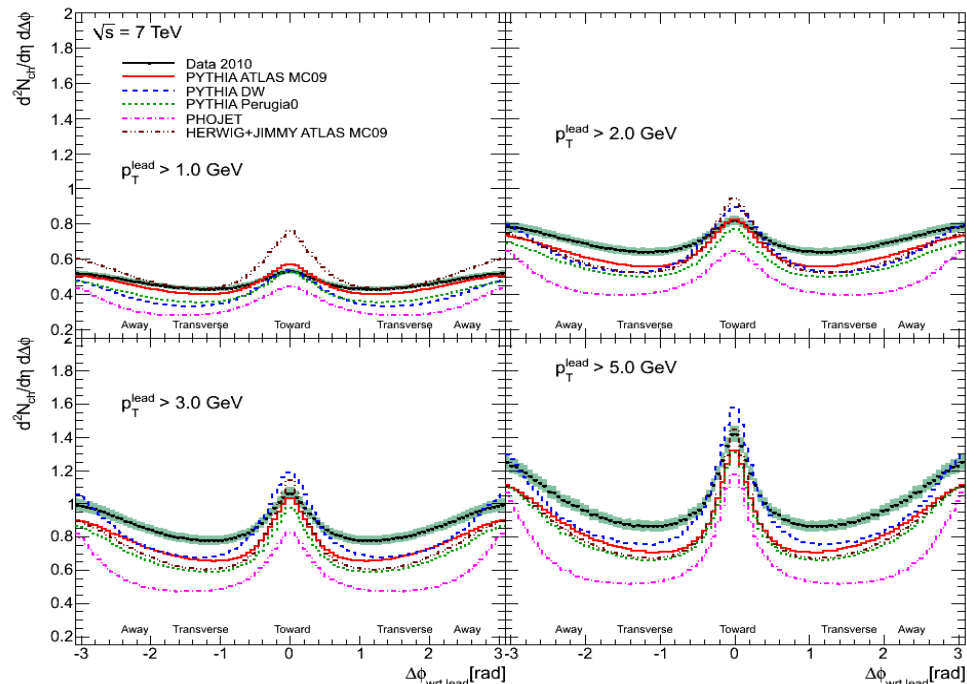
Charged particles



All particles

Comparing with the tracking results

charged particles



all particles

